El Ansari, Rokaya and Craze, Madeleine L. and Miligy, Islam and Diez-Rodriguez, Maria and Nolan, Christopher C. and Ellis, Ian O. and Rakha, Emad and Green, Andrew R. (2018) The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in Luminal B tumours. Breast Cancer Research . ISSN 1465-542X (In Press)

The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in Luminal B tumours

Rokaya El Ansari[a], Madeleine L. Craze[a], Islam Miligy[a], Maria Diez-Rodriguez[a], Christopher C. Nolan[a], Ian O. Ellis[a,b], Emad A. Rakha[a,b], Andrew R. Green[a]

[a]Academic Pathology, Division of Cancer and Stem Cells, School of Medicine, University of Nottingham, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB

[b]Breast Institute, Nottingham University Hospitals NHS Trust, Hucknall Road, Nottingham, NG5 1PB

Rokaya El Ansari: msxre2@nottingham.ac.uk

Madeleine L. Craze: Madeleine.Craze@nottingham.ac.uk

Miligy Islam: Islam.Abdelaziz@nottingham.ac.uk

Maria Diez-Rodriguez: Maria.Diez-Rodriguez@nottingham.ac.uk

Christopher C. Nolan: Chris.Nolan@nottingham.ac.uk

Ian O. Ellis: ian.ellis@nottingham.ac.uk

Emad A. Rakha: Emad.Rakha@nottingham.ac.uk

Andrew R. Green: andrew.green@nottingham.ac.uk

Correspondence:

Dr Andrew R. Green. Division of Cancer and Stem Cells, School of Medicine, University of Nottingham, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB

Tel: (44) 115 8231407, Email: andrew.green@nottingham.ac.uk

Abstract

Background: Breast cancer (BC) is a heterogeneous disease characterised by variant biology and patient outcome. The amino acid transporter, SLC7A5, plays a role in BC although its impact on patient outcome in different BC subtypes remains to be validated. This study aimed to determine whether the clinicopathological and prognostic value of SLC7A5 is different within the molecular classes of BC.

Methods: SLC7A5 was assessed at the genomic, using METABRIC data (n=1,980), and proteomic, using immunohistochemistry and TMA (n= 2664; 1,110 training and 1,554 validation sets), levels in well-characterised primary BC cohorts. SLC7A5 expression was correlated with clinicopathological and biological parameters, molecular subtypes, and patient outcome.

Results: SLC7A5 mRNA and protein expression were strongly correlated with larger tumour size, and higher grade. High expression was observed in triple negative (TN), HER2+, and luminal B subtypes. SLC7A5 mRNA and protein expression was significantly associated with the expression of the key regulator of tumour cell metabolism c-MYC, specifically in Luminal B tumours only (p=0.001). High expression of SLC7A5 mRNA and protein was associated with poor patient outcome (p <0.001) but only in the highly proliferative ER+/ luminal B (p=0.007) and HER2+ classes of BC (p=0.03). In multivariate analysis, SLC7A5 protein was independent risk factor for shorter Breast Cancer Specific Survival only in ER+ high proliferation tumours (p=0.02).

Conclusions: SLC7A5 appears to play a role in the aggressive highly proliferative ER+ subtype driven by *MYC* and could act as a potential therapeutic target. Functional assessment is necessary to reveal the specific role played by this transporter in the ER+highly proliferative subclass and HER2+ subclass of BC.

Key words: SLC7A5, breast cancer, prognosis

Introduction

Altered metabolic pathways have been readily accepted as part of the revised hallmarks of cancer where cancer cells are able to regulate their metabolism to provide energy and cellular building blocks required for their unremitting proliferation [1]. Many cancer cells are highly reliant on amino acids for their growth, not only because they are a nitrogen donor for the synthesis of nucleotides and other amino acids, but also because they activate mammalian target of rapamycin complex1 (mTORC1) through nutrient signalling pathways which in turn regulates protein translation and cell growth [2, 3]. There is also increasing evidence that oncogenes and/or tumour-suppressor genes can reprogram tumour cell metabolism including the direct regulation of the amino acid transporter, solute carrier family 7 member 5 (SLC7A5), by the oncogene *MYC* [4, 5] and the regulation of the glutamine transporter, SLC1A5, expression by the tumour suppressor Retinoblastoma (*Rb*) [6].

SLC7A5 is a sodium-independent transporter and acts as an amino acid exchanger by transporting large neutral amino acids such as leucine, phenylalanine and tryptophan by exchange with intracellular glutamine [7]. It therefore functions in supplying amino acids to cancer cells as well as maintaining intra-cellular leucine which is considered a master regulator of the mTORC1 signalling pathway [8-10]. For functional expression on the plasma membrane, SLC7A5 must heterodimerise with the heavy chain of SLC3A2 [7, 11].

It has been reported that SLC7A5 is highly expressed in a variety of cancers including oesophageal carcinoma [12], oral cancer [13] and lung adenocarcinoma [14]. SLC7A5 is co-expressed with the glutamine transporter, SLC1A5, in many cancers suggesting a functional coupling of these transporters in supporting tumour progression [15]. In this study, we aimed

to assess SLC7A5 gene copy number and mRNA expression, alongside protein expression in large and well-characterised annotated cohorts of BC to determine its biological, clinicopathological and prognostic value in the different BC molecular classes with particular interest in the highly proliferative aggressive subgroups.

Material and methods

SLC7A5 copy number and gene expression

A cohort of 1,980 BC tumours in the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) [16] was used to evaluate *SLC7A5* gene copy number aberrations (CNA) and gene expression. DNA/RNA was isolated from fresh frozen samples and genomic and transcriptional profiling were obtained using the Affymetrix SNP 6.0 and Illumina HT-12v3 platforms respectively. CNA were considered at the gene level by segments and Šidák correction [17] was applied for multiple testing. Gene expression data was pre-processed and normalised as described previously [16]. In this cohort, patients included were ER+ and /or Lymph Node negative (LN-) did not receive adjuvant chemotherapy, whereas ER- and LN+ patients received adjuvant treatment. X-tile (version 3.6.1, Yale University, USA) was applied to dichotomise SLC7A5 mRNA expression, based on prediction of Breast Cancer Specific Survival (BCSS). The association between the SLC7A5 mRNA expression and clinicopathological parameters, molecular subtypes, and patient outcome was investigated.

The online dataset, Breast Cancer Gene Expression Miner v4.0 (http://bcgenex.centregauducheau.fr) and the breast cancer TCGA data [18] were used for external validation of *SLC7A5* copy number/ or mRNA expression.

Patients and tumours

This study evaluated well-characterised cohorts of early stage primary operable invasive BC patients, who presented aged ≤ 70 years. Patients in the discovery set (n=1110) presented at Nottingham City Hospital between 1989 and 1998, while the validation set (n=1554) includes

patients who were presented between 1998 and 2006. Patient management was uniform and based on tumour characteristics by Nottingham Prognostic Index (NPI) and hormone receptor status. Patients within the NPI excellent prognostic group (score ≤3.4) received no adjuvant therapy, but those patients with NPI > 3.4 received tamoxifen if ER-positive [± Goserelin (Zoladex) in case the patients were premenopausal]. Conversely, classical cyclophosphamide, methotrexate and 5-flurouracil (CMF) were used if the patients were ER negative and fit enough to receive chemotherapy. None of the patients in this study received neoadjuvant therapy. Clinical history, tumour characteristics, information on therapy and outcomes are prospectively maintained. Outcome data included development and time to distant metastasis (DM) and Breast Cancer Specific Survival (BCSS).

There was no difference in the distribution of clinicopathological parameters between the discovery and validation cohorts or the METABRIC series of patients (all correlation coefficient ≥0.91, all $p<0.0001$) (Supplementary Table 1).

Western Blotting

The antibody specificity of anti-SLC7A5 (EPR17573, Abcam, UK) was validated using Western blotting in human embryonic Kidney (HEK) 293T over expression lysate (Origene Technologies, Rockville, MD, USA) and MDA-MB-175 (Luminal B-like), T47D and MCF7 (Luminal A) [19] breast cancer lysate (American Type Culture Collection; Rockville, MD, USA). A dilution of 1:200 of the primary antibody and 1:1000 HRP conjugated (Dako) secondary antibodies were applied. 5% milk /PBS-Tween (0.1%) (Marvel Original Dried Skimmed Milk, Premier Food Groups Ltd., UK) was used for blocking. Mouse monoclonal anti-β-actin primary antibody was used as a marker of endogenously expressed control. SLC7A5 bands were visualised using Enhance Chemiluminescence (ECL) showing a single specific band at the correct predicted size (40 KDa) for the SLC7A5 protein.

Tissue arrays and Immunohistochemistry

The discovery set (n=1,110) were arrayed as previously described using a single 0.6mm core sampled from the periphery of each invasive tumour [20]. The validation set (n=1,554) were similarly arrayed using a TMA GrandMaster (3D Histech). Immunohistochemical staining was performed on 4 μm TMA sections using Novolink polymer detection system (Leica Biosystems, RE7150-K). Briefly, tissue slides were deparaffinised with xylene and rehydrated through 3 changes of alcohol. Heat-induced antigen epitope retrieval was performed in citrate buffer (pH 6.0) for 20 minutes using a microwave oven (Whirpool JT359 Jet Chef 1000W). Endogenous peroxidase activity was blocked by Peroxidase Block for 5 minutes. Slides were washed with Tris-Buffered Saline (TBS, pH 7.6), followed by application of Protein Block for 5 minutes. Following another TBS wash, sections were incubated with the primary SLC7A5 antibody diluted at 1:50 in Leica antibody diluent (RE AR9352, Lieca, Biosysytems, UK) overnight at 4°C. Slides were washed with TBS followed by incubation with Post Primary Block for 30 minutes followed by a TBS wash. Novolink polymer was applied for 30 minutes. 3,3'-diaminobenzidine (DAB) chromogen was applied for 5 minutes. Slides were counterstained with Novolink haematoxylin for 6 minutes, dehydrated and coverslipped.

Stained TMA sections were scored using high resolution digital images (NanoZoomer; Hamamatsu Photonics, Welwyn Garden City, UK), at x20 magnification. Evaluation of staining for SLC7A5 was based on a semi-quantitative assessment of cores' digital images using a modified histochemical score (H-score) which includes an assessment of both the intensity and the percentage of stained cells [21]. Staining intensity was assessed as follows 0, negative; 1, weak; 2, medium; 3, strong and the percentage of the positively stained tumour cells was estimated subjectively. The final H-score was calculated multiplying the percentage of positive cells (0-100) by the intensity (0-3), producing a total range of 0-300.

Dichotomisation of protein expression was determined using x-tile software in predicting BCSS.

Immunhistochemical staining and dichotomisation of the other biomarkers included in this study were as per previous publications [20, 22-30]. ER and PgR positivity was defined as ≥1% staining. Immunoreactivity of HER2 in TMA cores was scored using standard HercepTest guidelines (Dako). Chromogenic in situ Hybridisation (CISH) was used to quantify HER2 gene amplification in borderline cases using the HER2 FISH pharmDx™ plus HER2 CISH pharmDx™ kit (Dako) and was assessed according to the American Society of Clinical Oncology guidelines. BC molecular subtypes were defined, based on tumour IHC profile and the Elston-Ellis [31] mitotic score as: ER+/HER2- Low Proliferation (mitotic score 1), ER+/HER2- High Proliferation (mitotic score 2 and 3), HER2-positive class: HER2+ regardless of ER status, Triple Negative: ER-, PgR- and HER2- [32]. Basal-like phenotype was defined as those tumours expressing cytokeratin (Ck) 5/6, and/or Ck14 and/or Ck17.

Statistical analysis

Statistical analysis was performed using SPSS 22.0 statistical software (SPSS Inc., Chicago, IL, USA). Spearman's correlation coefficient was carried out to examine the association between continuous variables. The chi-square test was performed for inter-relationships between categorical variables. For the continuous variables, differences between 3 or more groups were assessed using one way analysis of variance (ANOVA) with post-hoc Tukey multiple comparison test (for normalised data) or Kruskal-Wallis test (for non-normal distribution). Differences between two groups were assessed using t-test (normalised data) or Mann-Whitney test (non-normal distribution). Survival curves were analysed by Kaplan-Meier with Log Rank test. Cox's proportional hazard method was performed for multivariate analysis

to identify the independent prognostic/predictive factors. P-values were adjusted using Bonferroni correction for multiple testing. A p-value <0.05 was considered significant. The study endpoints were 10-year BCSS or Distant Metastasis Free Survival (DMFS).

Results

SLC7A5 genomic profiling

SLC7A5 was amplified in 0.3% and 0.6% of BC cases in the METABRIC and TCGA datasets, respectively, while deletion (deep and shallow) was detected in 56% and 68% of cases in the same cohorts respectively. Point mutations in *SLC7A5* were extremely rare, where TCGA data reported just one case with a missense mutation (Supplementary figure 4A) [33, 34]. *SLC7A5* is situated on chromosome 16 (16q24.2), all the annotated genes, which located on 16q [35] were selected to determine their CNV in relation to *SLC7A5* and assess whether these aberrations were locus-specific or involving large chromosomal segments. Both METABRIC and TCGA data showed a significant positive correlation between *SLC7A5* deletion and the deletion of all genes ($p<0.001$, Supplementary table 2). However, amplification of three genes (*FANCA*, *CBFA2T3* and *CDT1*) showed significant association with the amplified *SLC7A5* ($p\leq0.03$, data not shown) in the above-mentioned datasets together.

SLC7A5 expression in breast cancer

SLC7A5 protein expression was observed, predominantly in the membrane of invasive breast cancer cells, with expression levels varying from absent to high (Figure 1B and 1C). The distribution of the SLC7A5 protein expression was unimodal and left-skewed. The SLC7A5 mRNA expression had a normal distribution. Expression of SLC7A5 mRNA and protein were dichotomised using cut points derived from prediction of patient survival using X-tile (http://tissuearray.org; Yale University). Positive SLC7A5 expression (>15 H-score) was observed in 191/1,110 (17%) and 268/1,554 (17%) of cases in the discovery and validation sets respectively, while high SLC7A5 mRNA expression (log2 intensity >8) was observed in 1,019/1,923 (53%) of the METABRIC breast cancer cases. A total of 49/1,980 (2.4%) of cases

showed a copy number (CN) gain of *SLC7A5* and 530/1,980 (26.7%) showed a CN loss. A significant association was observed between *SLC7A5* copy number variation (CNV) and SLC7A5 mRNA expression ($p<0.001$, Figure 2).

SLC7A5 and clinicopathological parameters

Table 1 summarises the associations between SLC7A5 protein expression including larger tumour size, high tumour grade, and poor NPI (all $p<0.001$). Regarding BC metastatic sites, high SLC7A5 protein levels were associated with the development of distant metastases to the brain ($p<0.001$) and lung ($p=0.04$), while there was no association with developing DM to the bone or liver.

High SLC7A5 mRNA expression was significantly associated with larger tumour size (Figure 3A, $p<0.001$), higher grade (Figure 3B, $p<0.001$), positive nodal metastasis (Figure 3C, $p<0.001$) and poor NPI (Figure 3D, $p<0.001$). Both SLC7A5 mRNA and SLC7A5 protein were associated with medullary-like tumours. Where data was available, these associations were confirmed using the Breast Cancer Gene-Expression Miner v4.0 (Supplementary figure 1A, 1B) and the TCGA data (Supplementary figure 4B),. In addition *SLC7A5* copy number loss was significantly associated with good prognostic parameters including, lower grade and good NPI (Table 2, $p<0.001$). There was a positive association between *SLC7A5* copy number gain and *MYC* gain ($p<0.001$, Table 2).

SLC7A5 expression in molecular BC subtypes

SLC7A5 protein expression was associated with negative hormone receptor status and HER2+ tumours (all $p\leq0.002$, Table 3A) and it was highly expressed in Triple Negative (TN) and basal-

like phenotype malignancies compared to non-TN and non-basal-like tumours ($p<0.001$, Table 3A). Similarly, high expression of SLC7A5 mRNA was significantly associated with hormone receptor negative (ER- and PR-) and HER2+ tumours (all $p<0.001$, Table 3B). These results were in concordance with the findings of the Breast Cancer Gene-Expression Miner v4.0 (Supplementary Figure 1C-F) and TCGA data analysis (Supplementary figure 4C-4E).

When comparing the levels of SLC7A5 mRNA expression in the intrinsic (PAM50) subtypes [36], high expression was observed in Basal-like, HER2+ and Luminal B tumours (Figure 3F, $p<0.001$). Similarly, within the METABRIC Integrative Clusters, high SLC7A5 mRNA expression was associated with clusters 5 (ERBB2 amplified), 9 (Luminal B subgroup) and 10 (TN/basal-like) ($p<0.001$, Figure 3E). In the SCMGENE subtypes there was a higher expression of SLC7A5 mRNA in the ER+/HER2- high proliferation class (Luminal B) compared with the ER+/HER2- low proliferation class (Luminal A) ($p<0.001$, Figure 3G). Association of SLC7A5 mRNA with PAM50 subtypes was confirmed using the Breast Cancer Gene-Expression Miner v4.0 (Supplementary Figure 1G). Expression of SLC7A5 protein in the defined molecular subtypes showed a lower expression in the low proliferation tumours compared with the other subtypes ($p<0.001$, Table 1).

At the gene level, Basal-like showed a greater copy number gain of *SLC7A5* ($p<0.001$, Table 2), while *SCL7A5* copy number loss was mainly observed in the Luminal A subtype ($p<0.001$, Table 2).

SLC7A5 expression and other associated markers

The correlation of SLC7A5 mRNA with associated genes, using the METABRIC dataset, was investigated (Table 4). The genes were selected based on previous publications, being either a

regulatory genes or others that share or support the SLC7A5 biological function which focused mainly in glutamine transport and glutamine metabolism [2, 5, 15, 37-41]. There was a positive correlation between SLC7A5 mRNA expression and the expression of regulatory genes, several amino acid transporters and genes involved in the glutamine-proline regulatory axis. There was a positive relationship between SLC7A5 and MYC, mTOR and ATF4 ($p<0.001$) and the positive relationship between MYC, HIF2A and SLC7A5 was only observed in Luminal B tumours ($p=0.01$ and $p <0.001$ respectively).

High SLC7A5 mRNA expression was specifically associated with the enzymes involved with conversion of Glutamine (Gln) to proline, where PYCR1 and ALDH18A1 showed a positive relationship with SLC7A5 in almost all subtypes ($p<0.02$).

The majority of glutamine transporters were significantly associated with SLC7A5 expression primarily in triple negative tumours and to a lesser extent Luminal and HER2+ subtypes. SLC7A5 was significantly correlated with SLC1A5 in all subtypes ($p<0.02$).

TP53 mutations were also highly prevalent in breast tumours where there was high SLC7A5 mRNA expression ($p<0.001$, Table 3). Moreover, high SLC7A5 protein was positively associated with high p53 protein ($p<0.001$).

SLC7A5 protein expression was significantly expressed in breast tumours with high Ki67, and the upstream effector MYC ($p<0.001$, Table 5). The other investigated markers include the glutamine transporter SLC1A5 and other proteins involved in the glutamine metabolic pathways, as well as downstream signaling mTORC1. SLC1A5, GLS, PYCR1 and PIK3CA were significantly expressed in breast tumours with high expression of SLC7A5 ($p<0.001$), while the low expression was associated with high levels of p-mTORC1 ($p<0.001$) (Table 5).

SLC7A5 expression and patient outcome

Both high SLC7A5 mRNA ($p<0.001$) (Figure 4A) and protein ($p<0.001$) expression, in the discovery and validation sets, were associated with poor BCSS (Figure 5A, 5B). This association is anticipated as the cutoff was based on the prediction of BCSS.

While SLC7A5 mRNA expression was not predictive for BCSS in any specific molecular class (Figure 4B-4E), high expression of SLC7A5 protein was only predictive of shorter BCSS in ER+ high proliferation ($p=0.007$, Figure 5D) and HER2+ tumours ($p=0.03$, Figure 5F). There was no association between SLC7A5 protein and outcome in ER+ low proliferation (Figure 5C) or in TNBC (Figure 5E). In multivariate Cox regression analysis, SLC7A5 mRNA was a predictor of shorter BCSS independent of tumour size, grade, and lymph node stage ($p=0.006$, Supplementary table 3) but not in any specific subtype. However, SLC7A5 protein showed a significant result only in the ER+ high proliferation tumours ($p=0.02$, Table 6) and not with any other subtypes (data not shown).

Likewise, high SLC7A5 protein expression was associated with shorter distant metastases-free survival (DMFS) ($p<0.001$; Supplementary Figure 2A, 2B) within the ER+ high proliferation-class ($p=0.03$, Supplementary Figure 2D) but not with other subtypes (Supplementary Figure 2C, 2E, 2F). This association was found in the discovery set and validated in the validation set**.** The relationship between high SLC7A5 mRNA expression and poor patient outcome in ER+ disease, but not ER- disease, was confirmed using Breast Cancer Gene-Expression Miner (Supplementary figure 3A, 3B, 3C).

Discussion

Breast cancer is a heterogeneous disease with various subtypes [42] differing in terms of morphology, molecular and biological profiles, response to therapy and clinical behaviour. In addition, different subtypes exhibit disparity in their metabolic pathways and their nutritional needs. The most common form of BC (~55-80%) is ER+/luminal tumours [43, 44], and those that belong to this class are also variable in terms of recurrence, mortality rates and disease prognosis [43]. Therefore, understanding the biology of BC and exploring the metabolic pathways could help to identify potential novel therapeutic targets.

Cancer cells must alter their metabolism in order to satisfy the demands of necessary energy and cellular building blocks. It is widely known that amino acid transport systems play a principle role in sustaining the proliferation of cancer cells via supplying the required amino acids for protein synthesis as well as activation of nutrient signalling through the mTORC1 complex. This study has revealed, for the first time, that SLC7A5 is a key amino acid transporter in the more aggressive and highly proliferative ER+ tumours.

SLC7A5 is located in 16q24.2. This study showed that *SLC7A5* deletion, but not amplification, was significantly correlated with all the annotated genes located in the same chromosomal region, indicating that the deletion was not locus-specific. Interestingly, E-cadherin (*CDH1*), which located in 16q22.1, was also implicated. It is widely known that most lobular tumours harbour loss of heterozygosity (LOH) at chromosome 16 and miss the wild type *CDH1* allele [45]. In this study, approximately 40% of METABRIC cases, with *SLC7A5* loss, were associated with the invasive lobular histology. In addition, SLC7A5 protein expression in lobular carcinoma has a relatively lower mean rank value compared to the other histological

subtypes, confirming that deletions involve large segments of q16 which can reflect the BC phenotype.

SLC7A5 is widely expressed in many human cancers and various cancer cell lines [46]. The current study included two large discovery and validation cohorts to confirm the significant association between the high SLC7A5 protein expression and the poor prognostic clinico-pathological parameters, including larger tumour size, higher grade and poor NPI. Furthermore, high SLC7A5 expression was significantly associated with the higher expression of the proliferative marker (Ki67). This supports the results of previous studies which demonstrated that these two biomarkers are significantly correlated in tongue cancer [47], neuroendocrine carcinoma of the lung [48], thymic carcinomas [49] and breast cancer [50], indicating that SLC7A5 is critical for proliferation in cancer cells.

With respect to the breast cancer ER+ subtypes, SLC7A5 expression was lower in ER+ tumours that have low proliferation (luminal A subtype) compared with the highly proliferative ER+ (luminal B) malignancies, and it was primarily associated with poor patient outcome and shorter DMFS in the latter class only. This is most likely due to their heavier energy and nutrient requirements for cell survival, proliferation, and metastasis. This is anticipated, as it has been shown that over-expression of SLC7A5 is actively involved in the proliferation of vascular smooth muscle cells [51] and it is co-expressed with vascular endothelial growth factor (VEGF) in primary and metastatic sites of many cancers [37] which may be implicated for the metastatic process. In this study, luminal B subtype, showed the most significant positive correlation between the mRNA expression of SLC7A5 and VEGFA. In this regard, Bartlett et al included SLC7A5 as a part of the five gene Mammostrat® immunohistochemistry panel, where the higher expression is used to predict RFS, DRFS and OS in ER+ breast cancer during endocrine therapy [52]. However, they did not consider the different molecular subtypes of BC.

SLC7A5 mRNA and protein was also highly expressed in TNBC and HER2+, in concordance with Furuya et al [50]. However, in these subtypes the significant association between SLC7A5 protein expression and patient outcome was only observed in the HER2+ tumours. Among all BC subtypes, SLC7A5 protein expression was an independent predictor for short BCSS in ER+ high proliferation tumours only. In this regard, the larger sample size of ER+ high proliferation cases might reflect the stronger significance compared with the smaller sample size of HER2+ and TNBC tumours. We therefore suggest that further confirmation in larger cohorts of HER2+ and TN tumours is warranted.

Previous studies have showed regulation of SLC7A5 by other proteins including the tumour oncogene Myc which induces SLC7A5 [4, 5]. In the current study, the relationship between SLC7A5 and other regulatory proteins in both mRNA and protein expression was investigated. A positive relationship was observed between SLC7A5 and Myc in both protein and mRNA levels, and this correlation remained significant only in luminal B subtype, when different subtypes were investigated. ATF4-dependent transcripts also encodes for SLC7A5 upon amino acid deprivation [39] and this study observed a positive correlation between ATF4 and SLC7A5 gene expression in line with expectations. A previous study showed that activation of the HIF2α pathway increases mTORC1 activity by upregulating expression of the amino acid carrier SLC7A5 [38] and the current study confirmed the positive correlation between HIF2α and SLC7A5, which was only observed in luminal B tumours. SLC7A5 functions by importing essential amino acids to cancer cells and research has detailed the role of amino acids, particularly leucine, in activating mTORC1, which in turn controls protein translation, cell proliferation and prevent apoptosis in malignant cells [2, 3]. This study showed a positive correlation between SLC7A5 and mTOR at the mRNA level. However, the protein levels of SLC7A5 showed conflicting results, where high SLC7A5 expression was associated with the lower expression of the mTORC1 phosphorylated at ser (2448), which was included in this

study. This was unsurprising as Cheng et al confirmed that phosphorylation of mTORC1 at ser (2448), which is stimulated by growth factors, was mutually exclusive with mTORC1 phosphorylated at thr (2446), which is regulated by amino acids [53]. These observations may explain why SLC7A5 over-expression is primarily associated with poor outcome only in the high proliferative ER+ tumours.

This study further investigated the association of SLC7A5 expression with other glutamine transporters, in which some variability in the expression of amino acid transporters across molecular subtypes was observed. For example, TN subtype was the main class which showed association with the transporters required for glutamine influx. Perhaps because it depends on delivery of glutamine instead of synthesis. In contrast, a positive correlation between SLC7A5 and glutamine synthase enzyme GLUL was observed in Luminal A tumours, suggesting that this subtype might rely on glutamine neosynthesis rather than uptake. SLC1A5 functionally couples with SLC7A5 to allow the cellular influx and efflux of glutamine, as SLC1A5 mediates uptake of glutamine, while SLC7A5 uses intracellular glutamine concentrations to adjust essential amino acid cytoplasmic pool for metabolic demands and signalling to mTORC1 [15]. Here we observed that SLC7A5 and SLC1A5 are positively correlated in all the BC subtypes.

Previous studies have raised awareness and revealed the importance of the Proline-Glutamine (Pro-Gln) regulatory axis in BC. SLC7A5 appears to have a pivotal role in this regulatory axis, as its expression was highly associated with the enzymes which mediate glutamate degradation to form the amino acid proline, which has been shown to play a role in assisting tumour growth by different mechanisms [54].

Blocking SLC7A5 using its inhibitor, 2-aminobicyclo-(2,2,1)-heptane-2-carboxylic acid (BCH), efficiently decreased colony formation of MDA-MB-231 TNBC cells [55]. Even though the consequences of blocking SLC7A5 in the highly proliferative ER+ tumours remain

undetermined, this study suggests that SLC7A5 can potentially be used as a therapeutic target for luminal B BC**.**

Conclusion

This study revealed and confirmed that the glutamine transporter SLC7A5 was associated with poor prognostic characteristics and poor survival outcome. Over-expression of SLC7A5 appear to play a role in the proliferation and progression of the aggressive ER+ subtype of breast cancer, thus it could act as a potential therapeutic target. Functional assessment is necessary to reveal the specific role played by this amino acid transporter in the highly proliferative subclass and HER2+ BC.

List of abbreviations:

BC: Breast Cancer **METABRIC:** Molecular Taxonomy of Breast Cancer International Consortium **TN**: Triple Negative **HER2**: human epidermal growth factor 2 receptor **mTORC1**: Mammalian target of rapamycin complex1 **ER**: Estrogen Receptor **LN**: Lymph node **BCSS**: Breast Cancer Specific Survival **DM**: Distant Metastasis **ECL**: Enhance chemiluminescence **TMA**: Tissue microarray **PgR**: Progesterone Receptor **CISH**: Chromogenic in situ Hybridisation **CK**: Cytokeratin **DMFS**: Distant Metastasis Free Survival **CNV**: Copy Number Variation **NPI**: Nottingham Prognostic Index **Gln**: Glutamine **Pro**: Proline **VEGF**: Vascular Endothelial Growth Factor **BCH**: 2-aminobicyclo-(2,2,1)-heptane-2-carboxylic acid **CNA**: Copy Number Aberration **LOH**: Loss of Heterozygosity **CDH1**: E-cadherin.

Declarations:

Ethic Approval and consent to participate: This study was approved by the Nottingham Research Ethics Committee 2 under the title 'Development of a molecular genetic classification of breast cancer'.

Consent for publication: Not applicable

Availability of data and material: The datasets generated during the current study are available from the Corresponding Author on reasonable request.

Competing interests: Not Applicable

Funding: Not applicable

Authors' contribution:

RE contributed to writing, IHC staining, scoring, data analysis and interpretation; MLC contributed to writing and reviewing the manuscript; IM helped in scoring and double scoring; MDR and CCN helped in the laboratory work; IOE and EAR contributed to writing and reviewing the manuscript; ARG contributed to study design, data analysis and interpretation, writing and reviewing the manuscript.

Acknowledgments

We thank the Nottingham Health Science Biobank and Breast Cancer Now Tissue Bank for the provision of tissue samples.

Table 1. Clinicopathological associations of SLC7A5 protein expression in the discovery and validation breast cancer series.

	SLC7A5 protein							
	Discovery set				Validation set			
	n (%)	Mean	p-value	Adjusted p-value	n (%)	Mean	p-value	Adjusted p-value
Patient's Age								
≥50	395 (36)	582.20	0.003	**0.018**	469	859.46	$7.0x10^{-12}$	**<0.0001**
<50	714 (64)	539.95			1070	730.79		
Tumour size								
≥ 2.0cm	532 (48)	521.22	0.000002	**<0.0001**	939	729.47	$6.5x10^{-9}$	**<0.0001**
< 2.0cm	577 (52)	586.14			599	832.24		
Lympho-vascular Invasion								
Negative	722 (65)	545.17	0.138	0.55	1086	743.95	0.000007	**<0.0001**
Positive	382 (35)	566.35			451	829.31		
Site of distant metastasis								
Brain								
No	1044	546.62	0.0001	**0.0008**			N/A	
Yes	61 (6)	662.27						
Lung								
No	1003	547.12	0.007	**0.04**			N/A	
Yes	102 (9)	610.80						
Bone								
No	876 (79)	554.57	0.651	1.30			N/A	
Yes	229 (21)	547.00						
Liver								
No	949 (86)	551.43	0.568	1.70			N/A	
Yes	156 (14)	562.57						
	n (%)	**Mean**	**χ2**	**p-value**	**Adjusted p-**	**n (%)**	**Mean**	**χ2**	**p-value**	**Adjusted p-**
Tumour Grade										
1	190 (17)	450.77	171.5	$5.6x10^{-38}$	**<0.0001**	231	585.30	723.48	$7.7x10^{-72}$	**<0.0001**
2	366 (33)	473.19				622	647.46			
3	550 (50)	642.43				685	942.43			
Lymph Node Stage										
1	674 (61)	542.19	4.811	0.09	0.45	955	754.30	12.56	0.002	**0.004**
2	341 (31)	574.98				428	767.94			
3	91 (8)	556.75				153	858.69			
Nottingham Prognostic Index										

Good	332 (30)	458.21	102.4	$5.6x10^{-23}$	**<0.0001**	521	620.66	156.60	$9.8x10^{-35}$	**<0.0001**
Moderate	593 (53)	578.27				768	828.08			
Poor	184 (17)	654.64				246	892.49			
IHC Subtypes										
ER+/HER2- Low	250 (27)	391.92								
ER+/HER2- High	351(38)	419.54	178.4	$1.8x10^{-38}$	**<0.0001**			N/A		
Triple Negative	191 (20)	617.95								
HER2+	143 (15)	519.69								
Histological type										
Ductal (including	922 (83)	563.83				1335	782.19			
Lobular	101 (9)	454.17	69.05	$3.5x10^{-14}$	**<0.0001**	120 (8)	584.17	77.07	$7.2x10^{-16}$	**<0.0001**
Medullary	26 (2)	832.02				13	1257.27			
Miscellaneous	7 (0.6)	440.50				9 (0.6)	1037.78			
Special type	53 (5.4)	472.75				57	655.12			

Table 2. Copy number aberrations of *SLC7A5* in METABRIC breast cancer series and their associations with clinicopathological parameters, *MYC* copy number aberrations and breast cancer subtypes.

	SLC7A5 copy number							
	Gain				**Loss**			
	No	**Yes**	**χ2**	**Adjusted**	**No**	**Yes**	**χ2**	**Adjusted**
Age								
≥50	1520 (97.7)	36 (2.3)	0.405	1.572	1098 (70.6)	458 (29.4)	27.479	**<0.0001**
<50	372 (97.1)	11 (2.9)	(0.524)		321 (83.8)	62 (16.2)	($1.5x10^{-7}$)	
Tumour size								
≥ 2.0cm	1291 (97.0)	40 (3.0)	5.226	0.132	976 (73.3)	355 (26.7)	0.094	0.282
< 2.0cm	614 (98.7)	8 (1.3)	(0.022)		452 (72.7)	170 (27.3)	(0.759)	
Tumour Grade								
1	170 (100.0)	0 (0.0)	10.154		99 (58.2)	71 (41.8)	107.36	
2	756 (98.2)	14 (1.8)	(0.006)	**0.042**	495 (64.3)	275 (35.7)	$4.8x10^{-24}$	**<0.0001**
3	918 (96.4)	34 (3.6)			799 (83.9)	153 (16.1)		
Lymph Node Stage								
1	1012 (97.8)	23 (2.2)	0.474		726 (70.1)	309 (29.9)	10.425	
2	606 (97.4)	16 (2.6)	(0.789)	1.578	480 (77.2)	142 (22.8)	(0.005)	**0.02**
3	307 (97.2)	9 (2.8)			237 (75.0)	79 (25.0)		
Nottingham Prognostic Index								
Good	668 (98.2)	12 (1.8)	2.602		418 (61.5)	262 (38.5)	76.132	
Moderate	1071 (97.3)	30 (2.7)	(0.272)	0.080	864 (78.5)	237 (21.5)	($2.9x10^{-17}$)	**<0.0001**
Poor	192 (96.5)	7 (3.5)			168 (84.4)	31 (15.6)		
Histological type								
Ductal	1500 (97.2)	44 (2.8)			1154 (74.7)	390 (25.3)		
Lobular	145 (98.6)	2 (1.4)			88 (59.9)	59 (40.1)		
Medullary	30 (93.8)	2 (6.3)			30 (93.8)	2 (6.2)		
Miscellaneous	12 (100.0)	0 (0.0)	6.880		9 (75.0)	3 (25.0)	29.544	**0.0001**
Special type	113 (100.0)	0 (0.0)	(0.230)	1.150	74 (66.8)	39 (33.2)	(0.00001)	
PAM50 subtype								
Luminal A	710 (98.9)	8 (1.1)			423 (58.9)	295 (41.1)		
Luminal B	477 (97.7)	11 (2.3)			312 (63.9)	176(36.1)		
Basal	305 (92.7)	24 (7.3)			319 (97.0)	10 (3.0)		
HER2+	235 (97.9)	5 (2.1)	40.515	**<0.0001**	219 (91.3)	21 (8.7)	248.3	**<0.0001**
Normal-like	198 (99.5)	1 (0.5)	($3.3x10^{-8}$)		172 (11.9)	27(5.1)	($1.4x10^{-52}$)	
***MYC* Gain**								
No	1228 (98.9)	14 (1.1)	25.0	**<0.0001**	1446 (73.3)	528 (26.7)	0.132	1.432
Yes	703 (95.3)	35 (4.7)	($5.5x10^{-7}$)		4 (66.7)	2 (33.3)	(0.716)	

Table 3A. Association of SLC7A5 protein expression and the expression of other molecular biomarkers in the discovery and validation sets.

	SLC7A5 protein							
	Discovery set				Validation set			
	n (%)	Mean	p-value	Adjusted p-value	n (%)	Mean	p-value	Adjusted p-value
ER								
Negative	270 (25)	722.20	$3.2x10^{-48}$	**<0.0001**	300 (19)	1094.70	$4.6x10^{-76}$	**<0.0001**
Positive	827 (75)	492.45			1240 (81)	692.06		
PR								
Negative	435 (41)	619.12	$8.8x10^{-27}$	**<0.0001**	612 (42)	855.53	$1.3x10^{-34}$	**<0.0001**
Positive	630 (59)	473.54			853 (58)	645.09		
HER2								
Negative	921 (87)	521.78	0.00004	**0.0001**	1337 (92)	718.53	0.001	**0.002**
Positive	143 (13)	601.54			116 (8)	824.67		
Triple Negative								
No	896 (83)	503.50	$4.5x10^{-35}$	**<0.0001**	1286 (83)	696.76	$1.5x10^{-62}$	**<0.0001**
Yes	185 (17)	722.61			225 (17)	1094.6		
Basal Phenotype								
No	794 (74)	510.96	$6.8x10^{-13}$	**<0.0001**		N/A		
Yes	285 (26)	620.90						
P53 protein								
Negative	760 (72)	499.14	$4.1x10^{-13}$	**<0.0001**		N/A		
Positive	298 (28)	606.92						

Table 3B. Association of SLC7A5 mRNA expression and the expression of other molecular biomarkers in the METABRIC series.

	SLC7A5 mRNA expression				
	n (%)	**Mean**	**t-test**	**p-value**	**Adjusted p-value**
Estrogen Receptor					
Negative	474 (24)	9.543	26.90	$5.6x10^{-113}$	**<0.0001**
Positive	1506 (76)	7.943			
Progesterone Receptor					
Negative	940 (47)	8.862	18.73	$1.07x10^{-71}$	**<0.0001**
Positive	1040 (53)	7.841			
HER2					
Negative	1733 (88)	8.216	-12.35	$1.1x10^{-29}$	**<0.0001**
Positive	247 (12)	9.095			
Triple Negative (ER-, PR-, HER2-)					
No	1660 (84)	8.065	-22.12	$1.9x10^{-73}$	**<0.0001**
Yes	320 (16)	9.676			
Basal Phenotype					
No	1645 (83)	8.036	-25.70	$1.5x10^{-91}$	**<0.0001**
Yes	329 (17)	9.788			
***TP53* mutation**					
Wild-type	721 (88)	8.132	-7.47	$1.2x10^{-11}$	**<0.0001**
Mutation	99 (12)	9.148			

Table 4. Correlation of *SLC7A5* expression with the expression of other related genes in the METABRIC data.

	SLC7A5 mRNA expression									
	All cases (n=1,980)		**Luminal A** (n=368)		**Luminal B** (n=367)		**HER2+** (n=110)		**Triple negative** (n=150)	
	Correlation Coefficient (p-value)					**Adjusted p-value**				
Regulatory and other associated genes										
MYC	0.133 ($2.4x10^{-9}$)	**<0.0001**	0.012 (0.752)	4.145	0.155 (0.001)	**0.019**	0.066 (0.310)	4.650	0.103 (0.062)	0.434
mTOR	0.085 (0.0001)	**0.001**	-0.005 (0.904)	1.824	0.088 (0.052)	0.728	-0.023 (0.723)	5.784	0.067 (0.226)	1.130
VEGFA	0.352 ($6.4x10^{-59}$)	**<0.0001**	0.166 (0.000008)	**0.0002**	0.260 ($5.3x10^{-9}$)	**<0.0001**	0.269 (0.00002)	**0.0005**	0.244 (0.000008)	**0.0002**
HIF2A	-0.050 (0.028)	0.168	-0.023 (0.536)	4.896	0.215 (0.000002)	**<0.0001**	0.112 (0.083)	1.328	-0.282 ($1.8x10^{-7}$)	**<0.0001**
ATF4	0.159 ($1.0x10^{-12}$)	**<0.0001**	-0.029 (0.433)	5.100	0.057 (0.208)	2.080	0.143 (0.026)	0.468	0.108 (0.050)	0.450
Glutamine-proline regulatory axis										
GLS	0.048 (0.032)	0.192	0.008 (0.829)	3.428	0.055 (0.222)	1.998	-0.006 (0.927)	4.635	-0.115 (0.038)	0.456
ALDH4A1	-0.053 (0.019)	0.133	0.018 (0.638)	4.512	0.063 (0.163)	1.793	-0.028 (0.663)	5.967	-0.134 (0.015)	0.225
PRODH	0.004 (0.858)	1.716	-0.034 (0.369)	4.763	-0.032 (0.483)	2.415	0.037 (0.573)	5.73	0.030 (0.582)	1.746
PYCR1	0.32 ($1.5x10^{-50}$)	**<0.0001**	0.143 (0.0001)	**0.001**	0.253 ($1.3x10^{-8}$)	**<0.0001**	0.210 (0.001)	**0.024**	0.303 ($1.9x10^{-8}$)	**<0.0001**
ALDH18A1	0.222 ($1.6x10^{-23}$)	**<0.0001**	0.151 (0.00004)	**0.0008**	0.144 (0.001)	**0.018**	0.168 (0.009)	0.180	0.356 ($2.9x10^{-11}$)	**<0.0001**
GLUL	-0.18 ($3.3x10^{-16}$)	**<0.0001**	0.134 (0.0003)	**0.005**	0.008 (0.863)	1.726	-0.001 (0.991)	1.982	-0.122 (0.028)	0.392
GLUD1	-0.38 ($4.3x10^{-69}$)	**<0.0001**	-0.161 (0.00001)	**0.0002**	-0.237 ($1.1x10^{-7}$)	**<0.0001**	-0.148 (0.022)	0.418	-0.112 (0.042)	0.420
Glutamine/Glutamate transporters										
SLC1A5	0.29 ($4.5x10^{-41}$)	**<0.0001**	0.170 (0.000005)	**0.0001**	0.150 (0.001)	**0.017**	0.208 (0.001)	**0.023**	0.25 (0.000002)	**<0.0001**
SLC3A2	0.17 ($1.1x10^{-14}$)	**<0.0001**	0.067 (0.072)	3.780	0.193 (0.00001)	**0.0002**	0.184 (0.004)	0.084	0.158 (0.004)	0.064
SLC6A19	0.004 (0.869)	1.738	0.041 (0.273)	4.428	-0.008 (0.859)	2.577	0.047 (0.473)	5.676	-0.103 (0.061)	0.488
SLC7A6	0.362 ($2.7x10^{-62}$)	**<0.0001**	0.254 ($5.3x10^{-12}$)	**<0.0001**	0.33 ($2.0x10^{-14}$)	**<0.0001**	0.284 (0.000008)	**0.0002**	0.071 (0.201)	1.206
SLC7A7	0.19 ($4.6x10^{-19}$)	**<0.0001**	0.007 (0.857)	2.712	0.085 (0.061)	0.793	0.041 (0.530)	5.830	-0.22 (0.00005)	**0.0001**

SLC7A8	-0.42 (1.1×10^{-88})	**<0.0001**	-0.115 (0.002)	**0.034**	-0.103 (0.022)	0.352	-0.203 (0.002)	**0.044**	-0.40 (3.9×10^{-14})	**<0.0001**
SLC7A9	-0.068 (0.002)	0.01	0.025 (0.510)	4.82	0.044 (0.333)	2.331	-0.123 (0.056)	0.952	0.283 (1.8×10^{-7})	**<0.0001**
SLC38A1	-0.10 (0.000003)	**<0.0001**	-0.041 (0.270)	3.549	0.039 (0.391)	2.346	0.053 (0.413)	5.369	0.113 (0.041)	0.451
SLC38A2	-0.055 (0.015)	0.120	-0.074 (0.048)	1.05	-0.103 (0.022)	0.330	0.007 (0.917)	5.502	-0.119 (0.032)	0.416
SLC38A3	0.18 (8.3×10^{-17})	**<0.0001**	0.140 (0.0001)	**0.002**	0.046 (0.311)	2.488	0.003 (0.958)	2.874	0.196 (0.0003)	**0.005**
SLC38A5	0.011 (0.627)	2.574	-0.069 (0.066)	1.08	-0.077 (0.090)	1.080	-0.017 (0.793)	5.551	-0.017 (0.757)	1.514
SLC38A7	0.306 (3.8×10^{-44})	**<0.0001**	0.270 (2.0×10^{-13})	**<0.0001**	0.32 (1.2×10^{-13})	**<0.0001**	0.064 (0.324)	4.536	0.177 (0.001)	**0.017**
SLC38A8	0.023 (0.312)	1.560	-0.019 (0.612)	4.466	0.011 (0.801)	3.204	-0.006 (0.930)	3.720	-0.039 (0.482)	1.928

Table 5. Correlation between SLC7A5 protein expression and other biomarkers in the discovery set.

	SLC7A5 protein		
Biomarker	**Correlation coefficient**	**P-value**	**Adjusted p-value**
c-MYC	0.164	$8.2x10^{-7}$	**<0.0001**
Ki67	0.311	$1.1x10^{-21}$	**<0.0001**
P-mTORC1	-0.150	0.00001	**<0.0001**
PIK3CA	0.190	$3.4x10^{-7}$	**<0.0001**
SLC1A5	0.331	$1.1x10^{-25}$	**<0.0001**
GLUD1	0.053	0.09	0.180
GLS	0.371	$2.2x10^{-30}$	**<0.0001**
PYCR1	0.283	$1.07x10^{-16}$	**<0.0001**

Table 6. SLC7A5 protein expression and patient outcome in the combined, discovery and validation, sets in all cases and in ER-positive high proliferation tumours.

	SLC7A5 protein					
	All cases			ER+ high proliferation		
Parameter	**Hazard ratio (95% CI)**	**p-value**	**Adjusted p-value**	**Hazard ratio (95% CI)**	**p-value**	**Adjusted p-value**
SLC7A5	1.001 (1.000-1.003)	0.063	0.126	1.004 (1.001-1.006)	0.006	**0.024**
Lymph nodestage	2.060 (1.813-2.341)	$1.7x10^{-28}$	**<0.0001**	1.756 (1.427-2.161)	$1.04x10^{-7}$	**<0.0001**
Size	1.365 (1.111-1.678)	0.003	**0.009**	1.169 (0.838-1.632)	0.358	0.716
Grade	2.454 (2.023-2.977)	$1.8x10^{-20}$	**<0.0001**	1.756 (1.154-2.672)	0.009	**0.027**

Supplementary Table 1. Clinicopathological parameters of the METABRIC and Nottingham discovery and validation series.

	METABRIC series n (%)	**Nottingham discovery set n (%)**	**Nottingham Validation set n (%)**
Age			
≥ 50 years	1426 (78.6)	714 (61.3)	1070 (69.5)
< 50 years	424 (21.4)	395 (38.7)	469 (30.5)
Tumour size			
≥ 2cm	1337 (68.2)	577 (52.0)	599 (38.9)
<2cm	623 (31.8)	532 (48.0)	939 (61.1)
Grade			
1	169 (9.0)	190 (17.2)	231 (15.0)
2	770 (40.7)	366 (33.1)	622 (40.4)
3	952 (50.3)	550 (49.7)	685 (44.5)
Tumour type			
Ductal (including mixed)	1545 (83.6)	922 (83.1)	1335 (86.9)
Lobular	148 (8.0)	101 (9.1)	120 (7.8)
Medullary-like	32 (1.7)	26 (2.3)	13 (0.8)
Miscellaneous	12 (0.6)	7 (0.6)	9 (0.6)
Special type	113 (6.1)	53 (4.9)	60 (3.9)
Vascular Invasion			
Definite	Not available	382 (34.6)	451 (29.3)
Negative/Probable		722 (65.4)	1086 (70.7)
Lymph Node Stage			
1	1035 (52.5)	674 (70.0)	955 (62.2)
2	623 (31.5)	341 (30.8)	428 (27.9)
3	315 (16.0)	91 (8.2)	153 (10.0)
Follow-up Status			
Alive	1070 (55.7)	569 (51.3)	1110 (72.2)
Died from Breast Cancer	505 (26.3)	365 (33.0)	282 (18.3)
Died from other causes	345 (18.0)	175 (15.7)	146 (9.5)
Estrogen Receptor			
Negative	472 (23.8)	270 (24.6)	300 (19.5)
Positive	1508 (76.2)	827 (75.4)	1240 (80.5)
Progesterone Receptor			
Negative	938 (47.4)	435 (40.8)	612 (41.8)
Positive	1042 (52.6)	630 (59.2)	853 (58.2)
HER2 status			
Negative	1734 (87.5)	921 (86.6)	1376 (89.9)
Positive	246 (12.5)	143(13.4)	155 (10.1)

Supplementary table 2. List of genes whose copy number loss was significant associatedwith *SLC7A5* deletion in the METABRIC and TCGA data.

Gene symbol	Description	Chromosomal location	TCGA P-value	METABRIC P-value
SIAH1	siah E3 ubiquitin protein ligase 1	16q12.1	$1.5x10^{-3}$	$1.3x10^{-3}$
ABCC11	ATP binding cassette subfamily C member 11	16q12.1	$1.5x10^{-3}$	$1.3x10^{-3}$
CYLD	CYLD lysine 63 deubiquitinase	16q12.1	$1.5x10^{-3}$	$1.3x10^{-3}$
MMP2	matrix metallopeptidase 2	16q12.2	$1.2x10^{-4}$	$1.8x10^{-4}$
RBL2	RB transcriptional corepressor like 2	16q12.2	$6.1x10^{-3}$	$4.3x10^{-4}$
AMFR	autocrine motility factor receptor	16q13	$3.1x10^{-5}$	$1.8x10^{-4}$
MMP15	matrix metallopeptidase 15	16q21	$1.2x10^{-4}$	$1.8x10^{-4}$
USB1	U6 snRNA biogenesis phosphodiesterase 1	16q21	$1.2x10^{-4}$	$1.8x10^{-4}$
CX3CL1	C-X3-C motif chemokine ligand 1	16q21	$1.2x10^{-4}$	$1.8x10^{-4}$
NOL3	nucleolar protein 3	16q22.1	$8.8x10^{-7}$	$3.3x10^{-3}$
NQO1	NAD(P)H quinone dehydrogenase 1	16q22.1	$2.3x10^{-8}$	$2.2x10^{-4}$
TERF2	telomeric repeat binding factor 2	16q22.1	$8.8x10^{-7}$	$8.02x10^{-3}$
CBFB	core-binding factor beta subunit	16q22.1	$8.8x10^{-7}$	$6.1x10^{-3}$
CDH1	cadherin 1	16q22.1	$1.2x10^{-5}$	$3.4x10^{-4}$
CDH3	cadherin 3	16q22.1	$4.3x10^{-6}$	$1.4x10^{-4}$
CTCF	CCCTC-binding factor	16q22.1	$8.8x10^{-7}$	$1.3x10^{-5}$
E2F4	E2F transcription factor 4	16q22.1	$8.8x10^{-7}$	$9.4x10^{-6}$
PHLPP2	PH domain and leucine rich repeat protein phosphatase 2	16q22.2	$1.3x10^{-7}$	$2.2x10^{-6}$
ZFHX3	zinc finger homeobox 3	16q22.2-16q22.3	$1.3x10^{-7}$	$1.1x10^{-6}$
BCAR1	BCAR1, Cas family scaffolding protein	16q23.1	$3.8x10^{-9}$	$1.8x10^{-5}$
WWOX	WW domain containing oxidoreductase	16q23.1-16q23.2	$3.7x10^{-8}$	$6.3x10^{-6}$
ATMIN	ATM interactor	16q23.2	$1.2x10^{-11}$	$3.1x10^{-5}$
MAF	MAF bZIP transcription factor	16q23.2	$9.5x10^{-11}$	$4.5x10^{-7}$
OSGIN1	oxidative stress induced growth inhibitor 1	16q23.3	$3.8x10^{-11}$	$1.1x10^{-7}$
CDH13	cadherin 13	16q23.3	$2.1x10^{-12}$	$1.1x10^{-6}$
FOXF1	FOXF1 adjacent non-coding developmental regulatory RNA	16q24.1	$2.4x10^{-13}$	$1.1x10^{-7}$
WFDC1	WAP four-disulfide core domain 1	16q24.1	$2.1x10^{-12}$	$1.1x10^{-6}$
FBXO31	F-box protein 31	16q24.2	$2.4x10^{-13}$	$1.1x10^{-7}$
FANCA	Fanconi anemia complementation group A	16q24.3	$2.1x10^{-12}$	$5.4x10^{-4}$
CBFA2T3	CBFA2/RUNX1 translocation partner 3	16q24.3	$2.1x10^{-12}$	$7.8x10^{-5}$
CDT1	chromatin licensing and DNA replication factor 1	16q24.3	$2.1x10^{-12}$	$7.8x10^{-5}$

Supplementary table 3. SLC7A5 mRNA and patient outcome.

	SLC7A5 mRNA		
Parameter	**Hazard ratio (95% CI)**	**p-value**	**Adjusted p-value**
SLC7A5	1.121 (1.039-1.209)	0.003	**0.006**
LN stage	1.877 (1.656-2.129)	$8.9x10^{-23}$	**<0.0001**
Size	1.550 (1.219-1.971)	0.0003	**0.001**
Grade	1.331 (1.108-1.599)	0.002	**0.006**

Figure legends

Figure 1: SLC7A5 protein expression in invasive breast cancer cores. A) Negative IHC expression, B) positive IHC expression.

Figure 2: *SLC7A5* copy number aberrations and relationship with mRNA expression in the METABRIC cohort using One-way ANOVA with post-hoc tukey test.

Figure 3: SLC7A5 mRNA expression and its association with clinicopathological parameters and molecular subtypes: A) SLC7A5 and tumour size, B) SLC7A5 and tumour grade, C) SLC7A5 and lymph node stage, D) SLC7A5 and NPI, E) SLC7A5 and METABRIC Integrative Clusters, F) SLC7A5 and PAM50 subtypes, G) SLC7A5 and SMCGENE subtypes in the METABRIC cohort using One-way ANOVA with post-hoc tukey test.

Figure 4. SLC7A5 mRNA and breast cancer patient outcome. A) SLC7A5 vs BCSS in all cases, B) SLC7A5 vs BCSS in Luminal A tumours, C) SLC7A5 vs BCSS in Luminal B tumours, D) SLC7A5 vs BCSS in Triple Negative tumours, E) SLC7A5 vs BCSS in HER2+ tumours.

Figure 5. SLC7A5 and breast cancer patient outcome. A) SLC7A5 vs BCSS in all discovery set cases , B) SLC7A5 vs BCSS in all validation set cases, C) SLC7A5 vs BCSS of ER+- Low Proliferation tumours in the combined discovery and validation cases, D) SLC7A5 vs BCSS of ER+- High Proliferation tumours in the combined discovery and validation cases, E) SLC7A5 vs BCSS of Triple Negative tumours in the combined discovery and validation cases, F) SLC7A5 vs BCSS HER2+ tumours in the combined discovery and validation cases.

Supplementary Figure 1: SLC7A5 gene expression and its association, using Breast Cancer Gene-Expression Miner v4.0, with: A) tumour grade, B) NPI, C) ER status, D) PR status, E) HER2 status, F) Triple Negative status, G) PAM50 subtypes.

Supplementary Figure 2. SLC7A5 vs DMFS in A) all cases in the discovery set, B) All cases in the validation set, C) ER+- Low Proliferation tumours in the combined discovery and validation set cases, D) ER+- High Proliferation tumours in the combined discovery and validation set cases, E) Triple negative tumours in the combined discovery and validation set cases, F) HER2+ tumours in the combined discovery and validation set cases.

Supplementary Figure 3: SLC7A5 mRNA and breast cancer patient outcome using Breast Cancer Gene-Expression Miner: K) Unselected cases, L) ER+ disease, M) ER- disease.

Supplementary Figure 4. SLC7A5 mRNA expression, in the TCGA data, and its association with A) Copy Number Alteration, B) staging system, C) ER status, D) PR status, E) HER2 status.

References

1. Hanahan D, and Robert A. Weinberg: **Hallmarks of cancer: the next generation**. *cell* 2011, **144**(5):646-674.

2. Bond P: **Regulation of mTORC1 by growth factors, energy status, amino acids and mechanical stimuli at a glance**. *Journal of the International Society of Sports Nutrition* 2016, **13**:8.
3. Bar-Peled L, Sabatini DM: **Regulation of mTORC1 by amino acids**. *Trends in cell biology* 2014, **24**(7):400-406.
4. Gao P, Tchernyshyov I, Chang TC, Lee YS, Kita K, Ochi T, Zeller KI, De Marzo AM, Van Eyk JE, Mendell JT *et al*: **c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism**. *Nature* 2009, **458**(7239):762-765.
5. Hayashi K, Jutabha P, Endou H, Anzai N: **c-Myc is crucial for the expression of LAT1 in MIA Paca-2 human pancreatic cancer cells**. *Oncology reports* 2012, **28**(3):862-866.
6. Reynolds MR, Lane AN, Robertson B, Kemp S, Liu Y, Hill BG, Dean DC, Clem BF: **Control of glutamine metabolism by the tumor suppressor Rb**. *Oncogene* 2014, **33**(5):556-566.
7. Kanai Y, Segawa H, Miyamoto K, Uchino H, Takeda E, Endou H: **Expression cloning and characterization of a transporter for large neutral amino acids activated by the heavy chain of 4F2 antigen (CD98)**. *The Journal of biological chemistry* 1998, **273**(37):23629-23632.
8. Jewell JL, Kim YC, Russell RC, Yu FX, Park HW, Plouffe SW, Tagliabracci VS, Guan KL: **Metabolism. Differential regulation of mTORC1 by leucine and glutamine**. *Science (New York, NY)* 2015, **347**(6218):194-198.
9. Kimball SR: **Regulation of translation initiation by amino acids in eukaryotic cells**. *Progress in molecular and subcellular biology* 2001, **26**:155-184.
10. Bonfils G, Jaquenoud M, Bontron S, Ostrowicz C, Ungermann C, De Virgilio C: **Leucyl-tRNA synthetase controls TORC1 via the EGO complex**. *Molecular cell* 2012, **46**(1):105-110.
11. Yanagida O, Kanai Y, Chairoungdua A, Kim DK, Segawa H, Nii T, Cha SH, Matsuo H, Fukushima J, Fukasawa Y *et al*: **Human L-type amino acid transporter 1 (LAT1): characterization of function and expression in tumor cell lines**. *Biochimica et biophysica acta* 2001, **1514**(2):291-302.
12. Kobayashi H, Ishii Y, Takayama T: **Expression of L-type amino acid transporter 1 (LAT1) in esophageal carcinoma**. *Journal of surgical oncology* 2005, **90**(4):233-238.
13. Yoon JH, Kim IJ, Kim H, Kim HJ, Jeong MJ, Ahn SG, Kim SA, Lee CH, Choi BK, Kim JK *et al*: **Amino acid transport system L is differently expressed in human normal oral keratinocytes and human oral cancer cells**. *Cancer letters* 2005, **222**(2):237-245.
14. Nakanishi K, Matsuo H, Kanai Y, Endou H, Hiroi S, Tominaga S, Mukai M, Ikeda E, Ozeki Y, Aida S *et al*: **LAT1 expression in normal lung and in atypical adenomatous hyperplasia and adenocarcinoma of the lung**. *Virchows Archiv : an international journal of pathology* 2006, **448**(2):142-150.
15. Fuchs BC, Bode BP: **Amino acid transporters ASCT2 and LAT1 in cancer: partners in crime?** *Seminars in cancer biology* 2005, **15**(4):254-266.
16. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y *et al*: **The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups**. *Nature* 2012, **486**(7403):346-352.
17. Šidák Z: **Rectangular confidence regions for the means of multivariate normal distributions**. *Journal of the American Statistical Association* 1967, **62**(318):626-633.
18. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, Zhang H, McLellan M, Yau C, Kandoth C *et al*: **Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer**. *Cell* 2015, **163**(2):506-519.
19. Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, Kwei KA, Hernandez-Boussard T, Wang P, Gazdar AF *et al*: **Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery**. *PLoS One* 2009, **4**(7):e6146.
20. Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JF, Macmillan D, Blamey RW, Ellis IO: **High-throughput protein expression analysis using tissue microarray**

technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *International journal of cancer* 2005, **116**(3):340-350.
21. McCarty KS, Jr., McCarty KS, Sr.: **Histochemical approaches to steroid receptor analyses**. *Seminars in diagnostic pathology* 1984, **1**(4):297-308.
22. Green AR, Aleskandarany MA, Agarwal D, Elsheikh S, Nolan CC, Diez-Rodriguez M, Macmillan RD, Ball GR, Caldas C, Madhusudan S *et al*: **MYC functions are specific in biological subtypes of breast cancer and confers resistance to endocrine therapy in luminal tumours**. *Br J Cancer* 2016, **114**(8):917-928.
23. Aleskandarany MA, Rakha EA, Ahmed MA, Powe DG, Ellis IO, Green AR: **Clinicopathologic and molecular significance of phospho-Akt expression in early invasive breast cancer**. *Breast cancer research and treatment* 2011, **127**(2):407-416.
24. Abdel-Fatah TM, Powe DG, Ball G, Lopez-Garcia MA, Habashy HO, Green AR, Reis-Filho JS, Ellis IO: **Proposal for a modified grading system based on mitotic index and Bcl2 provides objective determination of clinical outcome for patients with breast cancer**. *The Journal of pathology* 2010, **222**(4):388-399.
25. Elsheikh S, Green AR, Aleskandarany MA, Grainge M, Paish CE, Lambros MB, Reis-Filho JS, Ellis IO: **CCND1 amplification and cyclin D1 expression in breast cancer and their relation with proteomic subgroups and patient outcome**. *Breast cancer research and treatment* 2008, **109**(2):325-335.
26. Aleskandarany MA, Rakha EA, Macmillan RD, Powe DG, Ellis IO, Green AR: **MIB1/Ki-67 labelling index can classify grade 2 breast cancer into two clinically distinct subgroups**. *Breast cancer research and treatment* 2011, **127**(3):591-599.
27. Jerjees DA, Negm OH, Alabdullah ML, Mirza S, Alkaabi M, Hameed MR, Abduljabbar R, Muftah A, Nolan CC, Green AR *et al*: **The mammalian target of rapamycin complex 1 (mTORC1) in breast cancer: the impact of oestrogen receptor and HER2 pathways**. *Breast cancer research and treatment* 2015, **150**(1):91-103.
28. Aleskandarany MA, Rakha EA, Ahmed MA, Powe DG, Paish EC, Macmillan RD, Ellis IO, Green AR: **PIK3CA expression in invasive breast cancer: a biomarker of poor prognosis**. *Breast cancer research and treatment* 2010, **122**(1):45-53.
29. Lancashire LJ, Powe DG, Reis-Filho JS, Rakha E, Lemetre C, Weigelt B, Abdel-Fatah TM, Green AR, Mukta R, Blamey R *et al*: **A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks**. *Breast cancer research and treatment* 2010, **120**(1):83-93.
30. Green AR, Caracappa D, Benhasouna AA, Alshareeda A, Nolan CC, Macmillan RD, Madhusudan S, Ellis IO, Rakha EA: **Biological and clinical significance of PARP1 protein expression in breast cancer**. *Breast cancer research and treatment* 2015, **149**(2):353-362.
31. Elston CW, Ellis IO: **Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. C. W. Elston & I. O. Ellis. Histopathology 1991; 19; 403-410**. *Histopathology* 2002, **41**(3a):151-152, discussion 152-153.
32. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, Zackrisson S, Cardoso F: **Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up**. *Annals of oncology : official journal of the European Society for Medical Oncology* 2015, **26 Suppl 5**:v8-30.
33. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E *et al*: **The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data**. *Cancer discovery* 2012, **2**(5):401-404.
34. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E *et al*: **Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal**. *Science signaling* 2013, **6**(269):pl1.

35. Huret JL, Ahmad M, Arsaban M, Bernheim A, Cigna J, Desangles F, Guignard JC, Jacquemot-Perbal MC, Labarussias M, Leberre V *et al*: **Atlas of genetics and cytogenetics in oncology and haematology in 2013**. *Nucleic acids research* 2013, **41**(Database issue):D920-924.
36. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z *et al*: **Supervised risk predictor of breast cancer based on intrinsic subtypes**. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009, **27**(8):1160-1167.
37. Kaira K, Oriuchi N, Imai H, Shimizu K, Yanagitani N, Sunaga N, Hisada T, Tanaka S, Ishizuka T, Kanai Y *et al*: **l-type amino acid transporter 1 and CD98 expression in primary and metastatic sites of human neoplasms**. *Cancer science* 2008, **99**(12):2380-2386.
38. Elorza A, Soro-Arnaiz I, Melendez-Rodriguez F, Rodriguez-Vaello V, Marsboom G, de Carcer G, Acosta-Iborra B, Albacete-Albacete L, Ordonez A, Serrano-Oviedo L *et al*: **HIF2alpha acts as an mTORC1 activator through the amino acid carrier SLC7A5**. *Molecular cell* 2012, **48**(5):681-691.
39. Adams CM: **Role of the transcription factor ATF4 in the anabolic actions of insulin and the anti-anabolic actions of glucocorticoids**. *The Journal of biological chemistry* 2007, **282**(23):16744-16753.
40. Craze ML, Cheung H, Jewa N, Coimbra ND, Soria D, El-Ansari R, Aleskandarany MA, Cheng KW, Diez-Rodriguez M, Nolan CC: **MYC regulation of Glutamine-Proline regulatory axis is key in Luminal B breast cancer**. *British journal of cancer* 2017.
41. El Ansari R, McIntyre A, Craze ML, Ellis IO, Rakha EA, Green AR: **Altered glutamine metabolism in breast cancer; subtype dependencies and alternative adaptations**. *Histopathology* 2018, **72**(2):183-190.
42. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA *et al*: **Molecular portraits of human breast tumours**. *Nature* 2000, **406**(6797):747-752.
43. Dawson SJ, Rueda OM, Aparicio S, Caldas C: **A new genome-driven integrated classification of breast cancer and its implications**. *The EMBO journal* 2013, **32**(5):617-628.
44. Rakha EA, El-Sayed ME, Green AR, Paish EC, Powe DG, Gee J, Nicholson RI, Lee AH, Robertson JF, Ellis IO: **Biologic and clinical characteristics of breast cancer with single hormone receptor positive phenotype**. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2007, **25**(30):4772-4778.
45. Berx G, Cleton-Jansen AM, Strumane K, de Leeuw WJ, Nollet F, van Roy F, Cornelisse C: **E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain**. *Oncogene* 1996, **13**(9):1919-1925.
46. Wang Q, Holst J: **L-type amino acid transport and cancer: targeting the mTORC1 pathway to inhibit neoplasia**. *Am J Cancer Res* 2015, **5**(4):1281-1294.
47. Toyoda M, Kaira K, Ohshima Y, Ishioka NS, Shino M, Sakakura K, Takayasu Y, Takahashi K, Tominaga H, Oriuchi N *et al*: **Prognostic significance of amino-acid transporter expression (LAT1, ASCT2, and xCT) in surgically resected tongue cancer**. *British journal of cancer* 2014, **110**(10):2506-2513.
48. Kaira K, Oriuchi N, Imai H, Shimizu K, Yanagitani N, Sunaga N, Hisada T, Kawashima O, Iijima H, Ishizuka T *et al*: **Expression of L-type amino acid transporter 1 (LAT1) in neuroendocrine tumors of the lung**. *Pathology, research and practice* 2008, **204**(8):553-561.
49. Kaira K, Oriuchi N, Imai H, Shimizu K, Yanagitani N, Sunaga N, Hisada T, Ishizuka T, Kanai Y, Endou H *et al*: **L-type amino acid transporter 1 (LAT1) is frequently expressed in thymic carcinomas but is absent in thymomas**. *Journal of surgical oncology* 2009, **99**(7):433-438.
50. Furuya M, Horiguchi J, Nakajima H, Kanai Y, Oyama T: **Correlation of L-type amino acid transporter 1 and CD98 expression with triple negative breast cancer prognosis**. *Cancer science* 2012, **103**(2):382-389.

51. Liu XM, Reyna SV, Ensenat D, Peyton KJ, Wang H, Schafer AI, Durante W: **Platelet-derived growth factor stimulates LAT1 gene expression in vascular smooth muscle: role in cell growth**. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2004, **18**(6):768-770.
52. Bartlett JM, Thomas J, Ross DT, Seitz RS, Ring BZ, Beck RA, Pedersen HC, Munro A, Kunkler IH, Campbell FM *et al*: **Mammostrat as a tool to stratify breast cancer patients at risk of recurrence during endocrine therapy**. *Breast cancer research : BCR* 2010, **12**(4):R47.
53. Cheng SW, Fryer LG, Carling D, Shepherd PR: **Thr2446 is a novel mammalian target of rapamycin (mTOR) phosphorylation site regulated by nutrient status**. *The Journal of biological chemistry* 2004, **279**(16):15719-15722.
54. Phang JM, Liu W, Hancock CN, Fischer JW: **Proline metabolism and cancer: emerging links to glutamine and collagen**. *Current opinion in clinical nutrition and metabolic care* 2015, **18**(1):71-77.
55. Liang Z, Cho HT, Williams L, Zhu A, Liang K, Huang K, Wu H, Jiang C, Hong S, Crowe R *et al*: **Potential Biomarker of L-type Amino Acid Transporter 1 in Breast Cancer Progression**. *Nuclear medicine and molecular imaging* 2011, **45**(2):93-102.